

Range Data Fusion for Accurate Surface Generation from Heterogeneous Range Scanners

Mahesh Kr. Singh¹, K. S. Venkatesh¹ and Ashish Dutta²

¹*Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, India*

²*Department of Mechanical Engineering, Indian Institute of Technology Kanpur, Kanpur, India*

Keywords: Gaussian Mixture Model, Laser Range Scanner, Kinect, RGB-D image, Delaunay Triangulation.

Abstract: In this paper, we present a new method for range data fusion from two heterogeneous range scanners for accurate surface modeling of rough and highly unstructured terrain. First, we present the segmentation of RGB-D images using the new framework of the GMM by employing the convex relaxation technique. After segmentation of RGB-D images, we transform both the range data to a common reference frame using PCA algorithm and apply the ICP algorithm to align both data in the reference frame. Based on a threshold criterion, we fuse the range data in such a way that the coarser regions are obtained from Kinect sensor and finer regions of plane are obtained from the Laser range sensor. After fusion, we apply Delaunay triangulation algorithm to generate the highly accurate surface model of the terrain. Finally, the experimental results show the robustness of the proposed approach.

1 INTRODUCTION

The multi-range sensor data fusion is the process of combining the range information from redundant and/or complementary sensors, to produce a complete and accurate description of the targeting region. The range data fusion has a special significance to generate the good quality surface. Nowadays, the generation of dense 3D representations of the environment has gained more attention. Some of the first work focused on the fusion of range data by making an implicit function (Wheeler et al., 1998) and then polygonizing it using the marching cubes algorithm for high resolution surface reconstruction. In (Trevor et al., 2012), the combination of 2D lines and 3D planes with a high level representation and easy to be annotated with semantic data have generated an accurate 3D map with its high level features. As discussed in (An et al., 2012), the authors have presented a fast incremental method of extracting planes using 2D lines from 3D point clouds acquired sequentially from a tilted LRF over mobile robot. In (Klöß et al., 2012), the authors have built the 3D surface element grid maps and present Monte Carlo localization with the probabilistic observation models for 2D and 3D sensors in this map. In (Newcombe et al., 2011), the authors have presented a new method for real-time 3D modeling of complex and arbitrary indoor scenes in

variable lighting conditions using a Kinect sensor and commodity graphics hardware. They have fused all the streamed depth data into a single global implicit surface model of the observed scene. In (Lai et al., 2011), the authors have presented a new method for RGB-D based object recognition and detection using color and depth information. In (Johnson and Manduchi, 2002), the authors have proposed a probabilistic rule for adaptive resolution integration of 3D data which has collected from multiple distributed sensors. In (Singh et al., 2014), the authors have proposed a new method for range data fusion from two heterogeneous range scanners. They have exploited the terrain characteristic (i.e. coarser and finer region) to fuse the range data and generated accurate 3D fused surface of the planner environment.

In this paper, first we present the segmentation of RGB images using the new framework of the Gaussian mixture model by applying the convex relaxation technique. After segmentation of RGB-images, we extract the corresponding location in Depth images by calibration RGB and depth images (Herrera et al., 2012). Now we are able to detect the finer location in Kinect frame. Also, we obtain the range data of the same environment from the Laser range scanner. Using the PCA algorithm, we transform both the range data into a common reference frame, and apply the ICP algorithm to align both range data. Based on a

threshold criterion, we have fused the range data in such a way that the coarser regions are obtained from Kinect sensor and finer regions of plane are obtained from the Laser range sensor. After fusion, we apply the Delaunay triangulation method to generate the highly accurate surface model of the terrain.

The remainder of this paper is organized as follows: Section II describes the proposed method in detail. Section III presents the experimental results. Finally, we conclude this paper in Section IV.

2 THE PROPOSED METHOD

The steps of the proposed method are described as follows:

2.1 Range Data Acquisition Systems

For the fusion of range data, we have used two heterogeneous range sensor, i.e. Laser range scanner and Microsoft Kinect. The figure 1(a) shows the Laser range scanner which is designed at our robotics lab. In the Laser scanner system, the Laser projects laser line on the plane and camera captures the laser line profile. When range scanner moves over the object surface, the camera acquires images of the distorted pattern which are reflected by the object surface. The height of the objects are obtained by taking into account the distortion of the laser light stripe caused by their shapes. The designed Laser range scanner gives accurate range measurements of the large angular field with angular resolution 0.1125° . The accuracy of the range scanner is approximately $\pm 2\text{-}3$ mm throughout its range. The major advantage of designed range scanner: it gives accurate result, very high angular resolution, no correspondence issue because the camera acquires the illuminated scene to obtain the dense 3D geometric information in a single exposure. The disadvantage is its high scanning time due to sequentially scan the terrain. The fig.1(b) shows the Kinect sensor that was introduced in Nov 2010 by Microsoft for the Xbox-360 video game system. The detail description of Microsoft Kinect is described in the papers (Zhang, 2012; Khoshelham and Elberink, 2012). In (Khoshelham and Elberink, 2012), the authors have investigated the accuracy and resolution of Kinect depth data for indoor mapping applications. They have presented that the random error of depth measurement increase quadratic-ally with increasing the distance from the sensor and it ranges from few millimeters up to 4 cm at the maximum range of the sensor. For the mapping application, the working range should be within 1-3 meter distance



Figure 1: (a) Laser Range sensor designed at our robotics lab (b) The Microsoft Kinect system.

from the sensor otherwise the quality of data is deteriorated by noise and low resolution.

2.2 RGB-D Segmentation

In this section, we describe the method to unsupervised segmentation of RGB-D images using new framework of GMM using by employing convex relaxation approach.

2.2.1 Gaussian Mixture Model and EM Algorithm

A Gaussian mixture model is a probabilistic model that presumes all the sample points are generated from a mixture of a fixed number of Gaussian distributions with unknown parameters. Therefore a Gaussian mixture model is a weighted sum of M Gaussian component densities of x which is a D -dimensional measurement vector as given by the equation,

$$p(x/\mu_i, \sigma_i) = \sum_{i=1}^N \omega_i g(x; \mu_i, \sigma_i^2) \quad (1)$$

where ω denote the mixture ratio, p is the Gaussian pdf parameterized by mean μ_i and variance σ^2 . Given data, the parameters $\Theta = \{\omega, \mu, \sigma^2\}$ can be efficiently estimated through maximum likelihood estimation (MLE) using the EM algorithm, then clusters will be obtained through estimated parameters. The log-likelihood function for GMM is given by (McLachlan and Krishnan, 2007)

$$\mathcal{L}(\Theta) = \int_{\Omega} \log \sum_{i=1}^N \frac{\omega_i}{\sqrt{2\pi\sigma_i}} \exp \left\{ -\frac{[f(x) - \mu_i]^2}{2\sigma_i^2} \right\} dx \quad (2)$$

we derive the conclusion from EM algorithm that the E-step and M-step can guarantee $\mathcal{L}(\Theta^{t+1}) \geq \mathcal{L}(\Theta^t)$ during the updating process (i.e. $t \rightarrow t+1$), which denotes the local convergence of the EM algorithm.

2.2.2 GMM using Convex Relaxation Approach

In this section, we present GMM algorithm using a convex relaxation approach. The optimization of logarithm and summation function like

equation (2) is difficult task because the operations of these functions are not non-commutative. However, the solution of these function could be achieved by adding a convex relaxation term. (\mathbf{u}, \mathbf{v}) represent the a vector valued function such as $\mathbf{v}(x) = (v_1(x), v_2(x), v_3(x), \dots, v_N(x))$ and $\Delta = \{\mathbf{v} | 0 < v_i < 1, \sum_{i=1}^N v_i = 1\}$. The symbol Δ is defined as the convex relaxation term for non binary vector space $(\mathbf{v}) = \{\mathbf{v} | v_i = \{0, 1\}, \sum_{i=1}^N v_i = 1\}$. For commutativity of log-sum function, we have used a deduction of convex analysis (Rockafellar, 1997)

Lemma 1. *The Log-sum Commutativity operations of given a function $\alpha_i(x) > 0$, for any function $\beta_i(x) > 0$, we have*

$$-\log \sum_{i=1}^N \alpha_i(x) \exp[-\beta_i(x)] = \min_{\mathbf{v}(x) \in \Delta} \left\{ \sum_{i=1}^N [\beta_i(x) - \alpha_i(x)] v_i(x) + \sum_{i=1}^N v_i(x) \log v_i(x) \right\} \quad (3)$$

Now we set $\frac{\omega_i}{\sqrt{2\pi}\sigma_i} = \alpha_i(x)$, $\frac{\|f(x) - \mu_i\|^2}{2\sigma_i^2} = \beta_i(x)$ and apply lemma in equation (2). The optimization problem becomes

$$\begin{aligned} \hat{\Theta} &= \arg \max_{\Theta} \mathcal{L}(\Theta) = -\arg \min_{\Theta} \mathcal{L}(\Theta) \\ &= \arg \min_{\Theta} \left\{ \int_{\Omega} \min_{\mathbf{v}(x) \in \Delta} \left\{ \sum_{i=1}^N [\beta_i(x) - \log \alpha_i(x)] v_i(x) + \sum_{i=1}^N v_i(x) \log v_i(x) \right\} dx \right\} \end{aligned} \quad (4)$$

Now we introduce a functional $\hat{\varepsilon}(\Theta, \mathbf{v})$ with two variables Θ, \mathbf{v} .

$$\begin{aligned} \hat{\varepsilon}(\Theta, \mathbf{v}) &= \int_{\Omega} \sum_{i=1}^N \left[\frac{\|f(x) - \mu_i\|^2}{2\sigma_i^2} - \log \frac{\omega_i}{\sqrt{2\pi}\sigma_i} \right] v_i(x) dx \\ &+ \int_{\Omega} \sum_{i=1}^N v_i(x) \log v_i(x) dx \end{aligned} \quad (5)$$

Then we compute the minimizer of $\hat{\varepsilon}(\Theta, \mathbf{v})$ via the following alternating algorithm:

$$\begin{aligned} \mathbf{v}^{t+1} &= \arg \min_{\mathbf{v} \in \xi} \hat{\varepsilon}(\Theta^t, \mathbf{v}) \\ \Theta^{t+1} &= \arg \min_{\Theta} \hat{\varepsilon}(\Theta, \mathbf{v}^{t+1}) \end{aligned} \quad (6)$$

Where $t = 1, 2, \dots$ is the iteration number and Θ^0 is an initial guess.

2.2.3 The Basic Model

Now we consider the non-uniform intensity problem because Kinect sensor capture the RGB-D images

of unstructured terrain, which can be mathematically modeled as:

$$f(x) = \gamma(x)g(x) \quad (7)$$

where $g(x)$ is the ground truth image, $f(x)$ refers the observed data and $\gamma(x)$ refers to a smooth varying bias field. From (Li et al., 2008), we have taken assumptions that the bias field is non-negative and smoothly varying. In the nearest neighborhood circle centered at x i.e. $\gamma(y) = \gamma(x)$, for all $y \in O_x$.

Here we describe method in (Li et al., 2008) with statistical interpretation. Let us first focus on the neighborhood O_y centered at y , all the intensity $f(y)$ within neighborhood O_y have same pfd $p(x)$ with the parameter $\mu_i, \sigma_i^2, \beta(y)$. If we deal with different contributions to the cost functional $\varepsilon(\Theta, \mathbf{v})$ in terms of distance to centering point, then we consider to add some weights for each pixel. We have taken the Gaussian function with std σ , $G_{\sigma}(x) \approx 0$ when $x \notin O_y$. Therefore, integral domain of O_y is expanded to whole domain Ω . i.e.,

$$\varepsilon_y(\Theta) = -\int_{\Omega} G_{\sigma}(y-x) \log \sum_{i=1}^N \frac{\omega_i}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{\|f(x) - \mu_i\|^2}{2\sigma_i^2} \right\} dx \quad (8)$$

For the desirable segmentation result, we have taken the global information and the total cost functional becomes

$$\varepsilon(\Theta) = \int_{\Omega} \varepsilon_y(\Theta) dy \quad (9)$$

The segmentation problem can be solve by the minimization problem

$$\Theta = \arg \min_{\Theta} \varepsilon(\Theta) \quad (10)$$

However, the above optimization problem is very difficult to solve due to the presence of the log-sum function. As a result, we construct the above cost function with two variables mention in previous section. According to lemma 1, the final data term becomes

$$\begin{aligned} \hat{\varepsilon}(\Theta, \mathbf{v}) &= \frac{1}{2} \int_{\Omega} \int_{\Omega} \sum_{i=1}^N G_{\sigma}(y-x) \left[\frac{\|f(x) - \mu_i\|^2}{2\sigma_i^2} - \log \frac{\omega_i}{\sqrt{2\pi}\sigma_i} \right] v_i(x) dx dy \\ &+ \int_{\Omega} \sum_{i=1}^N v_i(x) \log \omega_i(x) dx + \frac{1}{2} \int_{\Omega} \sum_{i=1}^N v_i(x) \log \sigma_i^2 dx \\ &+ \int_{\Omega} \log \beta(y) dy + \int_{\Omega} \sum_{i=1}^N v_i(x) \log v_i(x)^2 dx \end{aligned} \quad (11)$$

we have to minimize $\hat{\varepsilon}(\Theta, \mathbf{v})$ under the constraint $\mathbf{v} \in \Delta$ to get the optimized \mathbf{v}^*, Θ^* . Inspired by (Wang

et al., 2009), we have introduced another regularization term which is guarantee the precise close-form solution of \mathbf{v} . The regularization term define as

$$\mathcal{R}(\mathbf{v}) = \int_{\Omega} \sum_{i=1}^N v_i(x) e(x) \int_{N(x;\eta)} [1 - v_i(y)] dy dx \quad (12)$$

where e is the edge detector function, $N(x;\eta)$ give the neighborhood centered at x with radius η . In the paper, we take $e = \frac{1}{1 + \|\nabla G_{\sigma^*} f\|}$. Therefore combining the data term and regularization term, we introduce our new model as follows

$$(\Theta^*, \mathbf{v}^*) = \arg \min_{\Theta, \mathbf{v} \in \Delta} \mathcal{L} := \hat{\varepsilon}(\Theta, \mathbf{v}) + \lambda \mathcal{R}(\mathbf{v}) \quad (13)$$

where $\lambda > 0$ is a regularization parameter that controls the trade-off between these two function. The parameter set Θ becomes

$$\Theta = \{\omega_1 \dots \omega_N, \mu_1 \dots \mu_N, \sigma_1^2 \dots \sigma_N^2\} \cup \{\beta(y)\}_y \quad (14)$$

This new model is inherently different from classical GMM. Firstly, the classical GMM is very sensitive to noise and lacks in spatial smoothness constraint while the above new model has a controlling parameter that makes it robust to noise. Secondly, the data term in traditional GMM is only works well for images that have almost piece-wise constant and it can not handle the images with non-uniform intensity. However, the new model incorporates local bias function information, global intensity and edges information. Therefore, it works well on the images with non-uniform intensity. The minimizing eq. (15), starting from a given initial value Θ^0 such that

$$\begin{aligned} u^{t+1} &= \arg \min_{u \in \xi} \mathcal{L}(\Theta^t, v) \\ \Theta^{t+1} &= \arg \min_{\Theta} \mathcal{L}(\Theta, v^{t+1}) \end{aligned} \quad (15)$$

The stopping criteria of above proposed method is $\|\mathcal{L}_t - \mathcal{L}_{t+1}\|^2 < \delta \|\mathcal{L}_t\|^2$. In this way, we segment the RGB-images obtained from Kinect. Now our aim to correlate the segmented the RGB-images to their corresponding Depth images. For calibration of as discussed in (Herrera et al., 2012), the calibration of Depth and the color image pair is done using planar surface and a simple checkerboard pattern. The same calibration method is used to establish the relation between two range sensors. After calibration, we can easily locate the objects in the environment.

2.3 Data Fusion

The time cost of data acquisition from Laser range scanner is high, but it provides a very high quality range data i.e. 2 mm precision throughout the range.

On the other hand the random depth error in Kinect (Khoshelham and Elberink, 2012) increases with increasing distance from the sensor, but it is low cost, compact range sensor and very fast relative to designed Laser range sensor. The terrain characteristic is determine in terms of surface elevation. The finer regions of terrain are determined as the regions whose elevation is greater than 4 cm. Therefore, the threshold criterion for finer regions is determined in terms of elevation. We use Kinect to scan the coarser regions of terrain and the finer regions are scanned from a Laser range scanner. Since the range data obtain from both the range scanners are in different coordinate systems. Thus, it is necessary to transform both the range data into a common reference frame. We apply the Principal component analysis (PCA)(Jolliffe, 2005) to both range data sets that orthogonality transform the data set to the new coordinate system such that the largest variance of the data is defined as first coordinate (i.e. first principle component) and so on. This new coordinate system is defined as a common reference frame. Both range data are transformed into this reference coordinate system. With the help of segmented depth data, we find their corresponding points in the transformed reference frame. The finer detailed regions are identified in the reference frame. Now we apply the ICP algorithm (Elseberg et al., 2012) to align both the data set in this frame. This alignment of the two heterogeneous range data in reference frame is much faster than alignment of the range data in different coordinate system. Based on a threshold criterion, we fuse the finer region of data which is obtained from Laser range scanner to Kinect's coarser regions data. The fused range data eliminate the demerits of the range scanners by complementing each other. To generate the surface, we apply the Delaunay algorithm to the fused range data. In this way, we reconstruct the accurate, realistic surface of the terrain.

3 EXPERIMENTAL RESULTS

The proposed fusion method is tested on real world data by creating different types of environment in our lab. In the experiment, we have used the two heterogeneous Laser range scanner and Kinect shown in Fig (1). The purpose of the range data fusion is to generate the accurate, realistic, and a fast 3D surface of the terrain and eliminates the demerits of both range scanners by complementing each other. In the experiments, two range data sets of the same environment are obtained from both scanners (i.e. Laser range scanner and Kinect). First Kinect captures the RGB

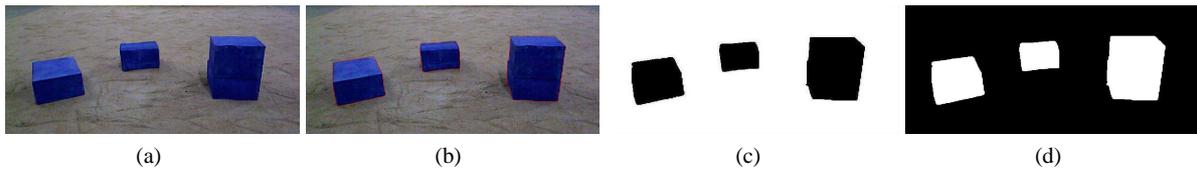


Figure 2: (a) Figure of different types of object placed in plane (b) Active contour map (c) Segmented blocks of RGB image and(d) Corresponding depth image.

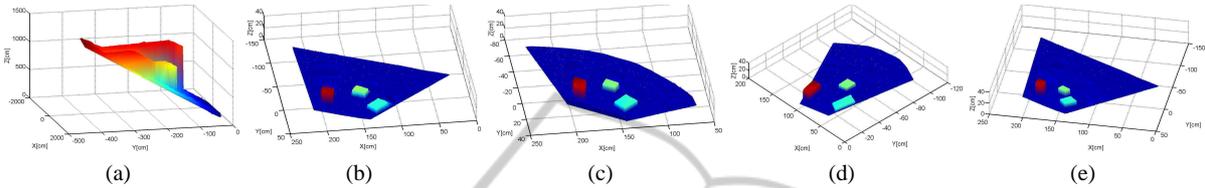


Figure 3: (a) The 3D surface of plane is obtained from Kinect (b) After applying ICP algorithm the aligned surface from Kinect and (c) from Laser range scanner (d) The Segmented fine region of surface from Laser range scanner (e) The accurately fused 3D surface model from both range sensors.

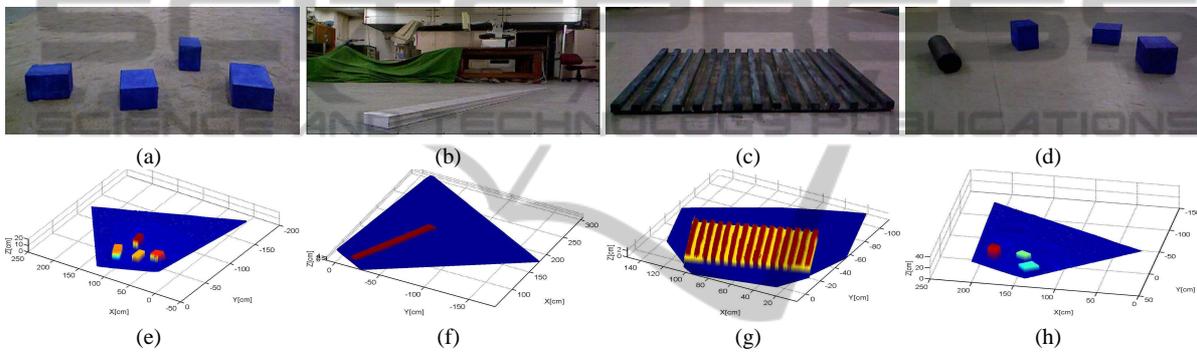


Figure 4: (a) Figure of different objects are placed on sandy plane (b) Figure of rectangular aluminum log (c) The plywood board consist of several stick's is placed (d) The different objects are placed (e) Shows the fused accurate 3D model of different objects (f) Shows the fused 3D model of aluminum log (g) Shows the fused 3D surface of plywood board (h) Shows the fused 3D surface of different objects.

and Depth image of the environment. The Fig 2(a) shows the RGB-image, The Fig 2(b) shows the active contour with new approach of GMM and Fig 2(c) shows the segmented objects in the image. To establish the relation between rgb image and depth image, we have used the calibration method given by (Herrera et al., 2012). The Fig 2(d) shows the segmented objects in the depth image. The finer regions of terrain are defined as the points of range data whose elevation is greater than 4 cm which is the threshold criterion. The range data acquired from both the range sensors are in different coordinate system. Therefore, it is necessary transform both range data into one common reference frame. We transform the both range data to a new coordinate system such that new set of uncorrelated variables axis, called principle components. The axis's of common reference frame are defined as the largest variance after transformation of data set to first coordinate system and so on. The figure 3(b-c) shows the aligned 3D surface of the terrain after applying the

ICP algorithm. The alignment of both range data in the common reference frame using ICP algorithm is much faster than directly apply the ICP algorithm for alignment of the range data in two different frames. We have defined the threshold for selecting fine region based on height data variation. Since Laser scanner time cost is high, therefore we have scanned the finer region of terrain using Laser scanner based on segmented depth data obtained from Kinect and rest regions are taken from Kinect i.e. we have retained the coarser detailed regions and erase the fine detailed region of Kinect range data. Now we have fused fine region range data acquired from Laser scanners to the coarser regions range data obtained from Kinect. Using Delaunay algorithm, we have generated surface of the terrain. The Fig. 3(e) shows the finally fused surface. In 3D fusion experiment, the relative sensor disparity of Laser range sensor relative to the Kinect in the reference frame is as the rotation matrix $R=[0.9704 \ 0.2418 \ 0;-0.2418 \ 0.9704 \ 0; 0 \ 0 \ 1.0000]$;

and translation vector $\mathbf{t}=[119.4117\ 127.9851\ 0]$. The alignment root mean square error of the fused data is approximately 3.2 mm. Many range data fusion experiments have been performed with different objects in the environment. Fig 4(a) shows the different kind of objects place in the plane and its accurate fused 3D terrain model is shown in Fig 4(e). Again the rectangular aluminum log is placed in plane as shown in Fig 4(b) and its fused terrain model is shown in Fig 4(f). Similarly, now we place the plywood board that has 15 rectangular log and different objects in the plane as Fig. 4(c-d). Fig. 4 (g-h) shows the accurately fused 3D model of the terrain. The resulting fused surface shows, the proposed method is applied to accurately, realistically and rapidly represent the real-world environment.

4 CONCLUSIONS

In this paper, we have presented a new approach for range data fusion from two heterogeneous range scanners (i.e. Laser range scanner and Microsoft Kinect) in order to integrate the merits of both scanners for the generation of accurate, realistic surface of the terrain. First, we have presented a new framework of the GMM using convex relaxation approach for segmentation of RGB-D images having inhomogeneous intensity. After transforming both the range data to common reference frame, we have applied the ICP algorithm to align these range data. The alignment method of two different range data in a common reference frame is much faster than directly apply the only ICP algorithm to their scanner coordinate system. In the fusion process, we have selected the coarser detailed region from Kinect and finer region from Laser scanner. The fused surface of the terrain is reconstructed using the Delaunay triangulation algorithm. In this way, we have generated a seamless integration of the terrain surface from two overlapping range data. The experimental results have shown the accurate 3D model of terrain from fused range data. The alignment rms error of the fused data has approximately 3-5 mm. So the main contribution of this paper is to present a range data fusion approach that eliminates the limitation of both the range sensors and generate the accurate surface modeling of rough and highly unstructured terrain.

REFERENCES

An, S.-Y., Lee, L.-K., and Oh, S.-Y. (2012). Fast incremental 3d plane extraction from a collection of 2d line

segments for 3d mapping. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 4530–4537. IEEE.

Elseberg, J., Magnenat, S., Siegwart, R., and Nüchter, A. (2012). Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration. *Journal of Software Engineering for Robotics*, 3(1):2–12.

Herrera, C., Kannala, J., Heikkilä, J., et al. (2012). Joint depth and color camera calibration with distortion correction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(10):2058–2064.

Johnson, A. E. and Manduchi, R. (2002). Probabilistic 3d data fusion for adaptive resolution surface generation. In *3D Data Processing Visualization and Transmission, International Symposium on*, pages 578–578. IEEE Computer Society.

Jolliffe, I. (2005). *Principal component analysis*. Wiley Online Library.

Khoshelham, K. and Elberink, S. O. (2012). Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors*, 12(2):1437–1454.

Kläß, J., Stückler, J., and Behnke, S. (2012). Efficient mobile robot navigation using 3d surfel grid maps. In *Robotics; Proceedings of ROBOTIK 2012; 7th German Conference on*, pages 1–4. VDE.

Lai, K., Bo, L., Ren, X., and Fox, D. (2011). A large-scale hierarchical multi-view rgb-d object dataset. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 1817–1824. IEEE.

Li, C., Kao, C.-Y., Gore, J. C., and Ding, Z. (2008). Minimization of region-scalable fitting energy for image segmentation. *Image Processing, IEEE Transactions on*, 17(10):1940–1949.

Newcombe, R. A., Davison, A. J., Izadi, S., Kohli, P., Hilliges, O., Shotton, J., Molyneaux, D., Hodges, S., Kim, D., and Fitzgibbon, A. (2011). Kinectfusion: Real-time dense surface mapping and tracking. In *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*, pages 127–136. IEEE.

Rockafellar, R. T. (1997). *Convex analysis*. Number 28. Princeton university press.

Singh, M. K., Venkatesh, K., and Dutta, A. (2014). Accurate 3d terrain modeling by range data fusion from two heterogeneous range scanners. In *India Conference (INDICON), 2014 Annual IEEE*, pages 1–6. IEEE.

Trevor, A. J., Rogers, J., and Christensen, H. I. (2012). Planar surface slam with 3d and 2d sensors. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 3041–3048. IEEE.

Wang, J., Ju, L., and Wang, X. (2009). An edge-weighted centroidal voronoi tessellation model for image segmentation. *Image Processing, IEEE Transactions on*, 18(8):1844–1858.

Wheeler, M. D., Sato, Y., and Ikeuchi, K. (1998). Consensus surfaces for modeling 3d objects from multiple range images. In *Computer Vision, 1998. Sixth International Conference on*, pages 917–924. IEEE.

Zhang, Z. (2012). Microsoft kinect sensor and its effect. *MultiMedia, IEEE*, 19(2):4–10.